U.S. Department of Energy Office of Science
Scientific Discovery through Advanced Computation Solicitation
National Collaboratories and High Performance Networks

Earth System Grid II
Turning Climate Datasets into Community Resources

Final Report

For the period June 1, 2001 – July 31, 2006

Principal Investigators (Laboratory)

Ian Foster
Mathematics and Computer Science Division, Argonne National Laboratory
Argonne, IL 60439
Tel: 630 252 4619 Fax: 630 252 5986
Email: foster@mcs.anl.gov

Dean N. Williams
Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory
Mail Stop L-264, 7000 East Ave, P.O. Box 808
Livermore, California 94550
Tel: 925 423 0145 Fax: 925 422 7675
Email: williams13@llnl.gov

Principal Investigator (University)

Don Middleton
Scientific Computing Division, National Center for Atmospheric Research
PO Box 3000; Boulder, Co. 80307-3000
Tel: 303 497 1250 Fax: 303 497 1804
Email: don@ucar.edu

Other Senior Personnel

Veronika Nefedova	Argonne National Laboratory
Alex Romosan	Lawrence Berkeley National Laboratory
Arie Shoshani	Lawrence Berkeley National Laboratory
Alex Sim	Lawrence Berkeley National Laboratory
Phil Jones	Los Alamos National Laboratory
Matthew Hecht	Los Alamos National Laboratory
Bob Drach	Lawrence Livermore National Laboratory
David Brown	National Center for Atmospheric Research
Luca Cinquni	National Center for Atmospheric Research
Peter Fox	National Center for Atmospheric Research
Jose Garcia	National Center for Atmospheric Research
Rob Markel	National Center for Atmospheric Research
Markus Stobbs	National Center for Atmospheric Research
Gary Strand	National Center for Atmospheric Research
Patrick West	National Center for Atmospheric Research
David Bernholdt	Oak Ridge National Laboratory
Mei-li Chen	Oak Ridge National Laboratory
Carl Kesselman	Information Sciences Institute, University of Southern California
Shishir S. Bharathi	Information Sciences Institute, University of Southern California
Ann Chervenak	Information Sciences Institute, University of Southern California

Table of Contents

Abstract

The Earth System Grid (ESG) II project, funded by the Department of Energy's Scientific Discovery through Advanced Computing program, has transformed climate data into community resources. ESG II has accomplished this goal by creating a virtual collaborative environment that links climate centers and users around the world to models and data via a computing Grid, which is based on the Department of Energy's supercomputing resources and the Internet. Our project's success stems from partnerships between climate researchers and computer scientists to advance basic and applied research in the terrestrial, atmospheric, and oceanic sciences. By interfacing with other climate science projects, we have learned that commonly used methods to manage and remotely distribute data among related groups lack infrastructure and under-utilize existing technologies. Knowledge and expertise gained from ESG II have helped the climate community plan strategies to manage a rapidly growing data environment more effectively. Moreover, approaches and technologies developed under the ESG project have impacted data-simulation integration in other disciplines, such as astrophysics, molecular biology and materials science.

A. Narrative

A.1. Overview

A.1.a. The Problem: Community Access to, and Analysis of, Climate Data

Science has become a data-intensive activity. New instrumentation, such as Earth observing satellites, space-based telescopes, and automated electron microscopes, are making detailed measurements and observations at a rate of 3 to 4 orders of magnitude greater than just five years ago. With unprecedented simulation capabilities, computational science is now also closely integrating models and observations in new ways.

Given high-resolution models and high-density observations, the repository of worldwide climate data is growing at an accelerated pace. The flood of simulations and observations has overwhelmed the climate community to such an extent that much of these data remain unexamined. In addition to the complexity and diversity of today's climate archives, the need to analyze both dataset ensembles and aggregations compound this problem, greatly hindering climate scientists from locating and sharing needed data.

To build a new and evolving infrastructure from scratch is a daunting task for climate scientists. Thus, the Earth System Grid (ESG) II project has approached cutting–edge computational technologies from the vantage point of developing hardware and software that will provide practical assistance to scientists in managing, sharing, and analyzing their data.

A.1.b. Project History

We took our first steps towards the realization of this vision in the project "Prototyping an Earth System Grid" (ESG I), which was initially funded under the Department of Energy's Next Generation Internet (NGI) program, with follow-on support from OBER and MICS. In this prototype, in the year 2000, we developed Data Grid technologies for managing the movement and replication of large datasets, and applied these technologies in a practical setting (i.e., an ESG-enabled data browser based on the PCMDI Climate Data Analysis Tools (CDAT)), achieving cross-country transfer rates of more than 500 Mb/s. Having demonstrated the potential for remotely accessing and analyzing climate data located at sites across the U.S., we won the "Hottest Infrastructure" award in the Network Challenge event.

While the ESG I prototype project substantiated a proof of concept ("Turning Climate Datasets into Community Resources"), the SciDAC Earth System Grid II project made this a reality. Our efforts targeted the development of metadata technologies (standard schema, XML metadata extraction based on netCDF, and a Metadata Catalog Service), security technologies (Web-based user registration and authentication, and community authorization), data transport technologies (GridFTP-enabled "OPeNDAP-G" for high-performance access, robust multiple file transport and integration with mass storage systems, and support for dataset aggregation and subsetting), as well as Web portal technologies to provide interactive access to climate data holdings. In 2003, a DOE-appointed external review committee praised these achievements, stating that "ESG is poised to make a substantial impact on the climate modeling community as well as many other data-intensive science disciplines."

In 2004, the National Center for Atmospheric Research (NCAR), a premier climate science laboratory, began to publish climate model data into ESG holdings. Late that same year, the Program for Climate Model Diagnosis and Intercomparison (PCMDI), an internationally recognized climate data center at Lawrence Livermore National Laboratory (LLNL), began its production service for climate model data germane to the Intergovernmental Panel on Climate Change (IPCC) 4th Assessment Report (AR4). (Because of international data requirements, restrictions, and timelines, the NCAR and PCMDI ESG data holdings were separated.) Since the release of these two institutional systems, ESG has become a world-renowned leader in developing technologies that provide scientists with virtual access to distributed data

and resources. To date, for example, some 4,000 users have downloaded more than 130 TB of data from these ESG sites.

A.1.c. The ESG II System: The Overall Architecture

The ESG II system can be described topologically as a Data Grid that connects important U.S. repositories of climate model data located presently at NCAR and at the Los Alamos (LANL), Lawrence Berkeley (LBNL), and Oak Ridge (ORNL) national laboratories (see figure 1). Data are stored either on disk farms for faster high-performance online access to frequently requested datasets (at LANL and NCAR), or on deep archives (National Energy Research Supercomputing Center (NERSC), NCAR Mass Storage System (MSS), and ORNL High Performance Storage System (HPSS)). , In addition, a dedicated ESG system has been implemented at LLNL to provide access to the IPCC AR4 data holdings, which also are stored online.

ESG II integrates a wide range of Grid and standard IT technologies with the goal of facilitating data publishing by climate modelers and data access by the worldwide climate community. The ESG web portal is the main access point to the system for both these data providers and users. The portal provides a central location for enforcing authentication, authorization and accounting (AAA) services, and brokers the formulation and submission of user data requests (transfer, download, sub-setting) among the distributed data nodes. The portal also provides the interface through which authorized data providers publish datasets into the system. All ESG II data nodes are instrumented with an RLS (Replica Location Services) server that indexes the physical files (or "replicas") available at that site. Online files are served via an LAHFS server (Lightweight Authorized HTTP File Server), while files on deep storage are requested and served via the SRM (Storage Resource Manager), which retrieves them from the archive and transfers them via GridFTP to a central disk cache, where they are made available by another LAHFS server. An OPeNDAP-G server also can be deployed at a data node to provide aggregation and sub-setting capabilities for online datasets. Finally, all the ESG II system components are continuously monitored, and users are notified whenever a service becomes unavailable.

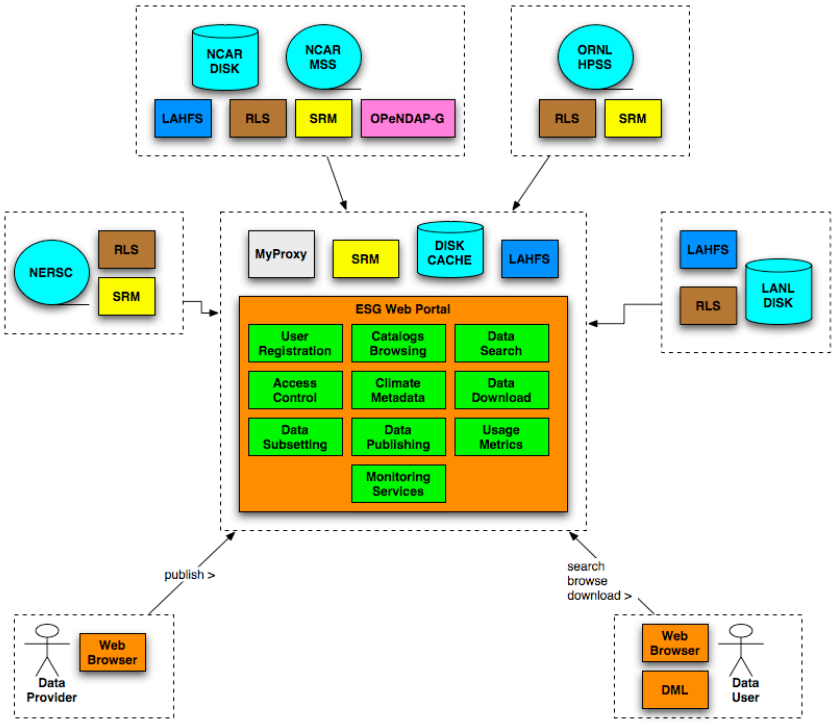

Figure 1: the ESG II topology

A.2. The ESG II Project Summary – Technology Integration and Innovation

A.2.a. Functional Specification for ESG II

- **Metadata:** The ESG II system offers unprecedented access to climate model data for the international scientific community. The overall system capabilities cover an extensive spectrum of functional areas pertinent to data and metadata management, access, and analysis. Recognized to be critical for the functionality and infrastructure of a Data Grid, two kinds of metadata are distinguished:
    - Descriptive or "logical" metadata include all the information necessary to accurately describe a climate model experiment, including scientific hypothesis, run input conditions, and output. These metadata are expressed in a Climate Model Metadata (CMM) schema specifically developed by ESG, and stored in a central relational database.
    - Location or "physical" metadata are used to access the data, and include information about possible multiple copies of the same data file. These metadata are stored in a system of cross-updating Replica Location Service (RLS) servers deployed at each ESG data node.

- **Data Discovery and Replication:** To provide data discovery for archived ESG files, we have deployed an RLS at five sites. Each RLS consists of a Local Replica Catalog (LRCs) that maintains mappings from logical file names to physical file locations at that site, and a Replica Location Index (RLI) that receives updates from all five LRCs in the system to track the local catalogs that contain mappings for a particular logical file. This system is used by the ESG portal to find the location of one or more physical files that match a user's request. Through the portal, the user can choose among available files for data access.

- **Data Publication:** In order to make data visible and accessible to end users, it first must be published into ESG. The Data Publication system scans files and directories, generates metadata in an XML markup language designed for ESG, and populates the RLS catalog and the metadata catalog. Additionally the publication system creates a collection of web pages that support the data portal browsing capability. Data can be published from local storage or remote mass store. Data on local disk can also be scanned to produce extra metadata of interest, such as the spatio-temporal extent and names of data variables. To date, over 250 TB of data have been registered in ESG catalogs.

- **Registration and Security:** The User Registration System was exclusively designed and developed for ESG use. It allows users with different permissions to access the ESG data holdings and services. To address the complex ESG authentication and authorization requirements, the standard Globus Toolkit Grid Security Infrastructure (GSI) was enhanced with advanced technologies, such as: MyProxy, and the Security Assertions Markup Language (SAML). SAML is the expression mechanism for authorization assertions.

- **Monitoring:** The climate community has come to depend on the ESG infrastructure as a critical resource. Failures of ESG components or services can severely disrupt the work of many scientists. To detect failures quickly and minimize infrastructure downtime, significant effort is spent on monitoring ESG resources. This monitoring uses the Globus Toolkit Monitoring and Discovery System (MDS4). The MDS4 Index Service collects information about ESG resources, while the MDS4 Trigger Service checks specified failure conditions and notifies system administrators when failures occur. We have used MDS for more than two years to monitor ESG resources and have significantly reduced the downtime of ESG infrastructure due to service failures.

- **Multi-site request execution:** Rather than painstakingly getting each file from its source location, ESG clients may wish to get a large number of files (regardless of their location on the Grid) in a single request (typically hundreds of files). Furthermore, it is impractical to expect users to learn how to interact with different file systems or mass storage systems at various sites. We have used Storage Resource Managers (SRMs) for this purpose. SRMs have been designed to provide a service of accepting multi-file requests, queuing each file request, getting the files from the source locations based on space availability (interacting with the various storage system involved), and streaming the files to the client. We have also used a tool for large-scale robust data replication tool, called DataMover, which interacts with SRMs for replicating thousands of files between mass storage systems. A client side version of this tool, called DataMover-Lite, was developed for use by clients to automate multi-file requests from SRMs into the clients' file systems.

- **Data Transport**: ESG datasets are large and are produced by model simulation codes designed for efficiency in output data handling and model-specific resource utilization, but not primarily with end-user applications in mind. As a result, ESG use-cases and user requirements implied a set of functional requirements for data transfers comprising single and multi-file transfer, data object creation and delivery, format handling, and virtual dataset creation in netCDF format with appropriate metadata. These ESG client uses required the ability to rapidly and reliably transfer large numbers of files from one location to another, extract and return portions of netCDF files, and assemble virtual datasets from a collection of files, while extracting user-specified variables.

- **Portal:** The ESG portal provides interactive, secure access to ESG services through the Internet using a web-based interface that is compatible with available web browsers. For scientists, the user interface enables the quick identification of data of interest through free-text searches and an intuitive browsing function, which drills into the data holdings following the structure of the ESG metadata schema. Once identified, data stored on disk can be downloaded directly through the portal via web hyperlink. For an archival storage site, the portal provides a data cart into which selected files can be placed. The data cart persists for the duration of a login session, so that a user can search or browse multiple files in different datasets, and then place them in the cart. A button on the datacart interface allows a single-click request to download the entire collection of selected files, with an e-mail message being sent to the user after all requested files have been collected from local or remote storage and moved to the portal's download cache, with a hyperlink supplied for each requested file. While active, data requests may be monitored through the portal to check for progress or errors. Aggregated datasets are available through an interface allowing extraction of data based on selected model variables and temporal, spatial and altitude parameters. For data providers, the portal provides a publishing interface for new data and the ability to update existing metadata. Metrics that detail the number of requests and total size of downloads per project or file also are viewable through the portal. Finally, the portal presents an interface allowing administrators to approve and reject account requests, renew account certificates, delete accounts, or modify user permissions via group and role affiliations.

A.2.b. Metadata

The critical importance of metadata to accurately describe, search, and validate climate model simulations was recognized at an early stage in the ESG project. We therefore spent considerable effort to create a metadata schema that would capture all information relevant to both simulation producers and users. The resulting schema is structured as a system of related objects, representing high-level concepts such as climate model 'project', 'ensemble', 'simulation', and open-ended hierarchies of 'datasets' containing 'datafiles'. Each object is described by a wide set of numerical and textual properties, some of which may be populated automatically, while others are filled (through the ESG web portal interface) by the domain

experts. These objects and properties were developed in strict collaboration with, and following the requirements of, the NCAR scientific community.

The ESG metadata schema was then mapped to a relational model so that the metadata contents could be stored in a relational database hosted on a centralized server. This database is populated during the ESG publishing process, and its content is used to create the metadata catalogs that users browse within the portal environment. It is also queried dynamically whenever a user wishes to display the full metadata pertaining to a specific logical object, such as the scientific hypothesis and input conditions for a given model simulation.

A.2.c. Grid Infrastructure Services

The need for data security posed a significant challenge for ESG at several levels. At the highest level, ESG nodes must function in the context of institutional computer/network security requirements and policies, while also interfacing with shared resources such as mass storage systems that are not operated exclusively by the ESG team. At a secondary level are those aspects of security that are in some way visible to ESG users (e.g., authentication and authorization). Finally, at a tertiary level are elements of the internal security model of the ESG system.

Our experience with "high-level" security issues is that the requirements of each institution are sufficiently different that they must be dealt with on a case-by-case basis. While the overall goals are fundamentally the same, the policies, procedures, and implementations at each site are sufficiently different that they cannot be dealt with collectively. Complicating our efforts at this level was the fact that computer security came under increasing scrutiny, and policies were correspondingly in flux, during the term of the ESG II project.

User-visible security measures respond primarily to the desire of the data owners to limit access to preliminary data (e.g., to allow their validation before general release). Because directly exposing X.509 certificates proved to be unpopular with users, we now implement password-based authentication in the portal (hiding the complexity of X.509 certificate management), coupled with group-based access controls which allow data owners to control group membership. Data delivery from HTTPS servers is secured by using a "Lightweight Authorized HTTP" scheme, which employs unique cryptographic tokens associated with the individual user and the specific files requested. When accessing the data on GridFTP servers, GSI authentication is used. The authorization statements (expressed in the form of SAML assertions) can be transferred to the server, either with GSI proxy certificates (as a non-critical extension) or via a GridFTP control channel.

A.2.d. Catalog Services

The ESG Metadata Catalog associates descriptive attributes with ESG datasets. The ontology for these attributes was defined by the Earth System Grid collaboration in the first years of the project. Users make metadata-based queries to this service to identify ESG datasets with certain characteristics--for example, datasets produced by a particular simulation model or that contain simulation results for a specific range of precipitation levels, surface temperatures, geographical areas, etc. Thus, users can query for related datasets of interest without needing to know specific filenames. We have experimented with different implementations for the Metadata Catalog over the course of the project, including the Globus Metadata Catalog Service and the OGSA Database Access and Integration Service. The current implementation uses a relational database deployed at NCAR that is queried by the portal; this implementation offers sufficient scalability to satisfy the metadata query load generated by users of the ESG portal.

Other catalogs deployed in ESG are the Globus Replica Location Services (RLS) deployed at five sites. RLS is a distributed data discovery service. It registers the location of data files at each site in a Local

Replica Catalog (LRC) by creating an association between a logical name for a file and its physical location on local storage. In addition, each LRC sends summaries of its state to higher-level Replica Location Index (RLI) nodes deployed at each site. By querying an RLI node, the ESG portal discovers all physical locations of a specified logical file. Since ESG datasets are typically very large, many of them are not replicated and reside only at a single site. Regardless of the amount of replication, RLS provides a scalable distributed mechanism for data registration and discovery. ESG also uses RLS attributes to store estimates of file size and other physical attributes that are useful in estimating the performance of data transfers.

During ESG Data Publication, publishers register new datasets in the Metadata Catalog by creating associations between the logical file names that comprise these datasets and the associated attribute values. Publishers also register the logical file names and physical file locations for new datasets in the RLS catalogs so that the portal can later discover the locations of these files.

A.2.e. Data Management Services

The transport of thousands of files is a tedious, error prone, but extremely important task in Climate modeling and analysis applications. The scientists who run models on powerful supercomputers often need the massive volume of data generated to be moved reliably to other sites. Often the source and destination storage systems are specialized mass storage systems, such as HPSS at NERSC and ORNL, and MSS at NCAR. During the analysis phase subsets of the data need to be accessed from various storage systems and moved reliably to a destination site where the analysis takes place. The automation of the file replication and subset extraction tasks require automatic space acquisition and reuse, and monitoring the progress of staging hundreds-to-thousands of files from the source mass storage system, transferring them over the network, and archiving them at the target mass storage system or transferring them to the clients. We have leveraged the software developed by another SciDAC supported middleware Storage Resource Manager (SRM) project to achieve robust multi-file transport for ESG. SRMs are software components that are placed in front of storage systems, accept multi-file requests, and manage the incremental access of the files from the storage system, and then use GridFTP to transport the files to their destination. Storage Resource Managers (SRMs) are Grid middleware components whose function is to provide dynamic space allocation and file management on shared storage components on the Grid. They are designed to provide effective sharing of files, by monitoring the activity of shared files, and making dynamic decisions on which files to replace when space is needed.

There are three important results achieved by ESG. The first achievement was motivated by a real practical need to Grid-enable the MSS at NCAR; that is, to be able to store files into and get files from the MSS at NCAR directly from remote sites. To achieve this goal, we have adapted a version of an SRM that was developed to work for HPSS to NCAR's MSS. The second achievement was motivated by requirements for data movement specified by ESG users, the most important of which is the ability to move entire directories in a single command. We have developed a software module, called the DataMover that can perform this task. Our work has resulted in a unique capability that effectively allows the equivalent of a Unix "cp –r" copy command across a heterogeneous online/archival storage environment, including NCAR's locally developed MSS. The third achievement was the development of a "Client-side DataMover" application, called DataMoverLite (DML) that works seamlessly with the web portal, allowing a user to make a selection of a large number of files and then triggering the launch of this application.

A.2.f. Monitoring

The ESG has many individual software components that run numerous systems at various institutions. It became clear early-on that in order to provide the required quality of service to users, it was necessary that all elements on the system be monitored on a regular basis for correct operation, and notifications sent to

project staff when anything required attention. We also found it useful to provide users with basic feedback on the operational status of the system, in order to help them understand why certain data might be temporarily unavailable or other problems that might arise.

The initial monitoring implementation was based on the Monitoring and Discovery System of the Globus Toolkit Version 3, and subsequently this has migrated to GT Version 4 components. Monitored ESG services include GridFTP servers, OPeNDAP, the ESG web portal, two HTTP data servers, RLS catalogs at five sites, SRMs at four sites, and three hierarchical mass storage systems. Monitored data are collected in an MDS4 Index service, which polls resource status at a configured frequency (currently every 10 minutes). The Index Service executes scripts that poll the status of resources. For example, when monitoring a GridFTP service, the script attempts to perform a small data transfer; for SRMs and their associated mass storage systems, the script runs an SRM client program that provides status. Resource information collected by the MDS4 Index service is queried by the ESG web portal to generate an overall picture of ESG infrastructure status that is displayed on the portal. Index service information is also polled periodically by the MDS4 Trigger service, which checks whether specified failure conditions are met. If so, it sends e-mail to ESG system administrators so they can quickly detect and address resource problems.

In addition, the RUDA resource usage-monitoring framework developed by the DOE Science Grid project was deployed in the ESG, and provided a successful demonstration of the technology. However, further work with RUDA was deferred until features such as server-side processing (now planned for the follow-on SciDAC project) make this kind of resource monitoring critical for ESG operation.

A.2.g. Aggregation and Sub-setting Services

Because OPeNDAP (Open-source Project for a Network Data Access Protocol) is a widely accepted concept, ESG reconfigured the architecture of the OPeNDAP core software and clients/servers to create OPeNDAP-G, a "Grid enabled" server that supports ESG security mechanisms and can utilize HTTP, HTTPS and GridFTP protocols for the data-transport mechanism.

In response to use cases and user requirements, ESG developed a seamless joining of data-extraction (subsetting) from different files and across distinct OPeNDAP servers. This "aggregation" was vital for the success of the ESG production release. In particular, performance metrics for generating aggregated datasets in the netCDF format (e.g. 200 MB in ~ 10 seconds) were used to drive design optimization for the OPeNDAP-G aggregation service. In addition, the assembly and transfer of virtual terabyte-scale datasets, was also a key factor in developing the data-transport design. As a result, ESG has built an end-to-end system connecting ESG clients (ESG portal, and the applications CDAT and NCL) to remote datasets over the HTTP and GridFTP protocols via the OPeNDAP-G client.

A.2.h. Application-specific Data Portals

Because of international data requirements, restrictions, and timeline, two separate web portals based on the same underlying integration of technologies were constructed to provide access to the vast data holdings made available through the ESG system.

The NCAR Data Portal serves data from a grid of distributed federated data centers (LANL, LBNL, NCAR, ORNL), comprising output and source code for the CCSM, PCM, CSIM, CLM and POP models, as well as supporting visualization and analysis software. The portal offers data providers and users a wide portfolio of functionality that includes searching and browsing of data catalogs, viewing of model-specific metadata, downloading of data files from either online storage or deep archives, extraction of specific subsets of virtually aggregated datasets, and publishing of new data into the system. Proper authentication and authorization is enforced for all operations, full logging of activities is maintained, and all system components deployed on the Grid are continuously monitored for availability of service. The NCAR portal

allows access to ~220 terabytes (TB) of distributed data and currently supports a community of almost 3,000 users, who collectively have downloaded ~25 TB to date.

The LLNL Data Portal focuses on the scientific needs of the 4th IPCC assessment and manages the related output from some 20 international climate models.

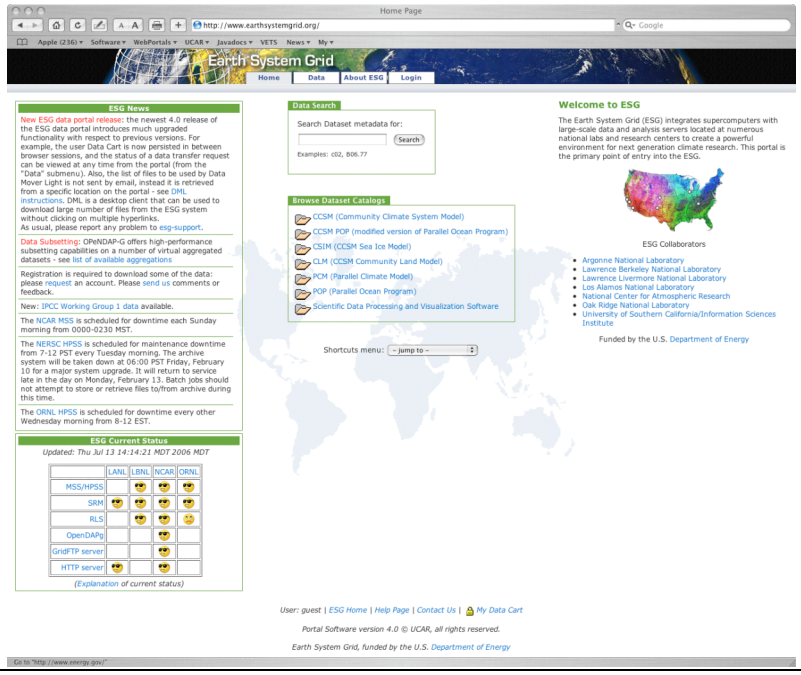

Figure 2: the ESG II web portal

A.3. Overall Impact of the Project

ESG now serves up a massive amount of data that are of critical importance to the climate community. In the beginning, we focused on publishing data generated by the Parallel Climate Model (PCM) and looked to additional data for inclusion. In particular, output data from the Climate Community System Model (CCSM) was published by NCAR and that from the Parallel Ocean Program (POP) model by LANL, while LLNL published multiple model data generated for the 2004 Intergovernmental Panel on Climate Change (IPCC) assessment. To support the infrastructural needs of the national and international climate community, ESG is providing crucial technology to securely access, monitor, catalog, transport, and distribute data in today's Grid computing environment.

A.3.a. International Panel on Climate Change Archive

In the autumn of 2004, ESG began its production service to distribute the Intergovernmental Panel on Climate Change (IPCC) / Working Group on Coupled Models (WGCM) data to the international climate community. The IPCC, which was jointly established by the World Meteorological Organisation (WMO) and the United Nations Environment Programme, carries out periodic assessments of the science of climate change. Fundamental to this effort is the production, collection and analysis of data from climate model simulations carried out by major international research centers. Analysis of a set of standard climate-change simulations from many modeling centers provides comprehensive understanding of the strengths and weaknesses of climate models. The IPCC and WGCM requested that PCMDI at LLNL collect model output data from these IPCC simulations, and distribute these to the community via ESG.

The ESG project has become internationally recognized for these efforts. Since December 2004, IPCC

model runs published to the climate community via the IPCC ESG portal totaled about 28 TB (66,200 files), and some 700 projects have registered to receive IPCC data for analysis. Overall, the number of files downloaded by the climate community totals about 460,000, which is equivalent to over 102 TB of data. The daily download rate averages about 200 GB. To date, the climate community has published some 250-research papers pertaining to analysis of the IPCC ESG data archive (see URL http://www-pcmdi.llnl.gov/ipcc/subproject_publications.php for citations).

A.3.b. CCSM and PCM Modeling Archive and Collaboration Sites

The Earth System Grid (ESG) project is meeting the critical challenge of effectively utilizing very large distributed simulation datasets. It has developed new data management and access systems that provide access to catalogs and data associated with the past six years of joint DOE/NSF climate modeling experiments using the Community Climate System Model (CCSM) and the Parallel Climate Model (PCM). Some 3,000 registered users now have ready access to nearly 220 TB of scientific data that are spread across several DOE supercomputer sites.

This user base represents climate scientists, analysts, educators, governments (both domestic and abroad), private industry, and many others. CCSM and PCM data accessed via ESG has been used in numerous scientific papers, impact analysis, urban planning, ecosystem monitoring, education, and other activities. In addition, the CCSM model code has been accessed from ESG over 400 times in the past year, enabling scientists, hardware and software engineers, universities and others to examine and learn how a state-of-the-art climate model works, and to provide suggestions and enhancements for its scientific accuracy, portability, and performance.

As of late Spring, the ESG team has been working closely with the CCSM community to publish CCSM model data automatically into the ESG archives. Collaborating with CCSM scientists and data providers, the ESG team is developing a web interface to the ESG metadata database that will allow the CCSM community to completely manage all information related to generating, defining, and archiving CCSM model simulation runs. This interface, still in beta stage, allows scientists to impose selective access control on the project runs, to sort information by any type, and to collaboratively enter data. The long-term goal is to tie the metadata ingestion process to the actual CCSM run workflow, so that model simulation metadata can be added automatically into the ESG data holdings.

A.3.c. POP Modeling Archive

The Climate, Ocean and Sea Ice Modeling (COSIM) project at Los Alamos National Laboratory (LANL) develops the Parallel Ocean Program (POP) and CICE model and uses these models to perform very high resolution simulations of the global ocean to examine the impacts of mesoscale eddies on the global circulation. LANL began participating in ESG to distribute these high-resolution datasets to its collaborators and the rest of the oceanographic community. While the numerous simulations and the total size of these data make their distribution difficult enough, their high-resolution characteristics present particular challenges owing to the size of each prognostic field (1.4 GB). In addition, the ocean data are on displaced-pole and tripole grids, making subsetting and other manipulation more complex. Finally, LANL has more stringent computer security requirements that those of other ESG collaborators, raising additional logistical barriers to the widespread distribution of these data.

To serve the ocean data, LANL acquired a dual-processor 64-bit Linux server with 5 TB of RAID disk which was placed on the laboratory's open network. Because this was the first 64-bit server in ESG, installation uncovered several bugs in the Grid software that took some time to fix. Once these issues were addressed, LANL began serving 100 GB worth of high-resolution North Atlantic Ocean data as a test case. Global high-resolution datasets now are being transferred to the server, and these will soon be published via the NCAR ESG portal, with NCAR processing the requests and accessing data from the LANL

archive. In the near future, we plan to configure the LANL server as a standalone ESG node, so as to be able to publish the data directly.

A.3.d. Specific Technology Impact

- **Registration System:** Because the registration system developed for ESG has been well accepted by the climate community, we decided to release this piece of technology to outside users. We developed a non-ESG specific application, PURSe (Portal-based User Registration System), which is now available for download. PURSe, a system for registering users of Web-based applications that use the Grid Security Infrastructure (based on PKI and X.509 certificates), coordinates the process of establishing Grid security certificates for new users when logging into an application website. Users who already have valid certificates can also easily register with portal applications. By leveraging the MyProxy certificate repository, PURSe shields web application users from the complexities of X.509 certificate management, thus enabling rapid registration of users.

- **Monitoring Infrastructure:** The monitoring requirements of ESG have been a main driver for improving the functionality of monitoring services in the Globus Toolkit over the last two years. From its initial use of Globus Toolkit Version 3 monitoring services, ESG has been one of the earliest adopters of these services, and has been instrumental in defining requirements for the GT Version 4 Monitoring and Discovery System. ESG requirements were particularly important in driving improvements in the functionality and scalability of the MDS4 Trigger Service. These MDS4 monitoring services are now available to all Globus users.

- **OPeNDAP-G:** Through our collaboration with the OPeNDAP community we propagated the improvements and augmentations made to the core OPeNDAP services for ESG back into the mainstream software releases that are now being made available to the broad data community as part of the OPeNDAP SERVER4 release. The modified (multi-tiered) architecture of the OPeNDAP core now provides for considerable flexibility in supporting a variety of data formats, client access requests, transport protocols, and security infrastructures.

- **NetCDF Markup Language:** The NetCDF Markup Language (NcML) was created from the joint collaboration of ESG and Unidata developers, with the initial original purpose of expressing aggregation information for virtual datasets spanning many individual data files comprising the full output stream of a specific model component. NcML is an XML encoding of the NetCDF data model, and can be used to fully represent a single data file, a collection of logically related files, and even a virtual dataset that defines additional properties not contained in the actual files. NcML has now become a community standard, and it is now actively supported and developed by Unidata.

A.3.e. A Vision for the Future: ESG-CET

ESG II, under SciDAC-1, brought major advances in climate modeling and the management and sharing of the distributed terascale data via Grid technology. Nevertheless, the climate research community still has a daunting task of locating, acquiring, and integrating collections of *both* simulation and experimental petascale data evolving multidisciplinary study of the Earth systems.

Building upon our previous success and experience, ESG's next challenge will be to establish the *Earth System Grid Center for Enabling Technologies* (ESG-CET), which is intended to resize ESG for the petascale realm and equip it to serve future scientific requirements. This environment will enable broad community access to, and deep analysis of, simulation and experimental data from a distributed network of sources.

Our high-level goals are driven by scientific objectives relevant to DOE scientific priorities over the next five years. In brief, these are:

- Sustain successful existing ESG services.
- Address scientific needs related to five-year projections of data management and analysis requirements, with a particular focus on:
    - Preparing for the IPCC 5th Assessment Report (AR5) in 2010.
    - Publishing and processing of the massive data produced by the *Climate Science Computational End Station (CCES)* at ORNL's NCCS/LCF.
    - Providing tools for climate model evaluation activities under the proposed SciDAC2 project, *A Scalable and Extensible Earth System Model for Climate Change Science.*

Figure 3 depicts the scientific data management and analysis requirements in relationship to the ESG development timeframe. Note that *a distributed testbed for IPCC AR5 must be in place by early 2009*.

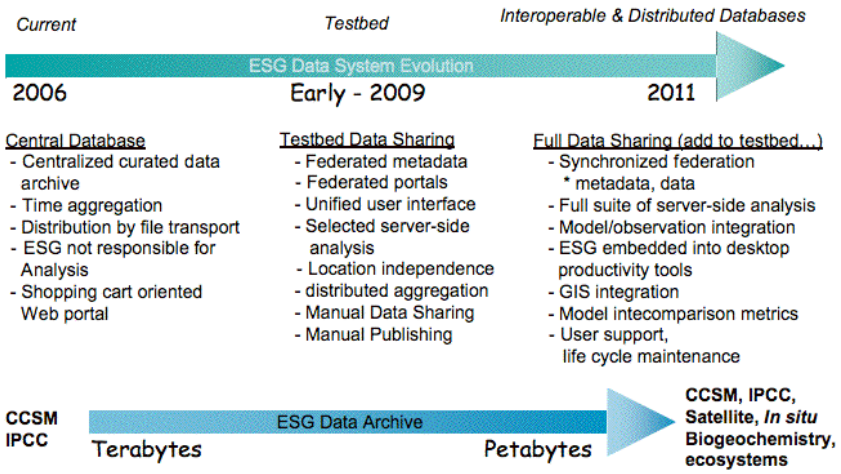

Figure 3: evolving ESG to the petascale: high-level ESG-CET roadmap

B. Appendix: ESG and Component Publications

1. Fox, P., Garcia, J. and West, P. OPeNDAP for the Earth System Grid, in preparation. Data Science Journal. 2006.
2. ESG. The Earth System Grid - Turning Climate Model Datasets Into Community Resources. http://www.earthsystemgrid.org, 2006.
3. IPCC. Intergovernmental Panel on Climate Change - 4th Assessment Report. http://www.ipcc.ch/activity/ar.htm, 2006.
4. IPCC. The Intergovernmental Panel on Climate Change (IPCC) publication list. http://www-pcmdi.llnl.gov/ipcc/subproject_publications.php, 2006.
5. PCMDI The LLNL/PCMDI ESG IPCC Portal. https://esg.llnl.gov:8443/index.jsp, 2006.
6. CCSM. CSSM - The Community Climate System Model. http://www.ccsm.ucar.edu/, 2006.
7. NCAR. LAHFS - A Lightweight Authorized HTTP File System. http://www.vets.ucar.edu/software_docs/public/NCAR_data_download/LAHFS.doc, 2006.
8. NCL. The NCAR Command Language (NCL). http://ngwww.ucar.edu/ncl/index.html, 2006.
9. LLNL/PCMDI. CDAT - Climate Data Analysis Tools. http://www-pcmdi.llnl.gov/software-portal/cdat/, 2006.
10. MyProxy, http://grid.ncsa.uiuc.edu/myproxy, 2005.
11. OPeNDAP: Open-source Project for a Network Data Access Protocol, www.opendap.org, 2005.
12. Portal-based User Registration Service (PURSE), www.grids-center.org/solutions/purse, 2005.
13. Pouchard, Line, Woolf, Andrew, and Bernholdt, David, Data Grid Discovery and Semantic Web Technologies for the Earth Sciences, International Journal on Digital Libraries. 5:2 (April 2005)
14. The Replica Location Service. http://www.globus.org/RLS, 2005.
15. Bernholdt, D., Bharathi, S., Brown, D., Chanchio, K., Chen, M., Chervenak, A., Cinquini, L., Drach, B., Foster, I., Fox, P., Garcia, J., Kesselman, C., Markel, R., Middleton, D., Nefedova, V., Pouchard, L., Shoshani, A., Sim, A., Strand, G. and Williams, D. The Earth System Grid: Supporting the Next Generation of Climate Modeling Research. Proceedings of the IEEE, 93 (3). 485-495. 2005.
16. Chervenak, A., R. Schuler, C. Kesselman, S. Koranda, B. Moe, Wide Area Data Replication for Scientific Collaborations. In Proceedings of the 6th IEEE/ACM Int'l Workshop on Grid Computing (Grid2005), (Seattle, WA, USA, 2005).
17. Pouchard, Line (Chair), Semantic Infrastructure for Grid Computing Applications Workshop, Fifth IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2005). Cardiff, Wales, UK. May 9-12, 2005.
18. Cai, M., A. Chervenak, M. Frank, A Peer-to-Peer Replica Location Service Based on a Distributed Hash Table. In Proceedings of the SC2004 Conference, (Pittsburgh, PA, 2004).
19. LIGO Project. Lightweight Data Replicator, http://www.lsc-group.phys.uwm.edu/LDR/. http://www.lsc-group.phys.uwm.edu/LDR/, 2004.
20. Storage Resource Management Project, http://sdm.lbl.gov/srm and http://sdm.lbl.gov/srm-wg, 2004.
21. Sim, A., Gu, J., Shoshani, A. and Natarajan, V., DataMover: Robust Terabyte-Scale Multi-file Replication over Wide-Area Networks. In Proceedings of the 16th International Conference on Scientific and Statistical Database Management (SSDBM), (2004).
22. Welch, V. Globus Toolkit Version 4 Grid Security Infrastructure: A Standards Perspective, http://www.globus.org/toolkit/docs/4.0/security/GT4-GSI-Overview.pdf, 2004.
23. Chawathe, Y., Ratnasamy, S., Breslau, L., Lanham, N. and Shenker, S., Making Gnutella-like P2P Systems Scalable. in ACM SIGCOMM 2003, (Karlsruhe, Germany, 2003).

24. Shoshani, A., Sim, A. and Gu, J. Storage Resource Managers: Essential Components for the Grid. in Nabrzyski, J., Schopf, J.M. and Weglarz, J. eds. Grid Resource Management: State of the Art and Future Trends, Kluwer Academic Publishers, 2003.
25. Pouchard, Line, Cinquini, Luca, Strand, Gary, et al., The Earth SystemGrid Discovery and Semantic Web Technologies, in Semantic WebTechnologies for Searching and Retrieving Scientific Data workshop. 2nd International Semantic Web Conference. Sanibel Island, Florida, October 20-24, 2003.
26. Pouchard, Line, Cinquini, Luca, Drach, Middleton, Don, Bernholdt, David, Chanchio, Kasidit, Chen, Meili, Foster, Ian, Nefedova, Veronika, Brown, David, Fox, Peter, Garcia, Jose, Strand, Gary, Williams, Dena, Chervenak, Ann, Kesselman, Carl, Shoshani, Arie, and Sim, Alex, An Ontology for Scientific Information in a Grid Environment: the Earth System Grid, in Agent-based Cluster and Grid Computing Workshop. Third IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2003). Tokyo, Japan, May 12-15, 2003.
27. Domenico, B., Caron, J., Davis, E., Kambic, R. and Nativi, S. Thematic Real-time Environmental Distributed Data Services (THREDDS): Incorporating Interactive Analysis Tools into NSDL. Journal of Digital Information, 2 (4). 2002.
28. Security Association Markup Language (SAML) Specification v.1.0, OASIS, http://www.oasis-open.org/committees/security/, 2002.
29. Globus Project The GridFTP Protocol and Software. 2002.
30. Globus Project. Grid Information Services / MDS, http://www.globus.org/mds/, 2002.
31. Stoica, I., Robert Morris, David Karger, M. Frans Kaashoek, and Hari Balakrishnan, Chord: A scalable peer-to-peer lookup service for internet applications. in ACM SIGCOMM, (2001).
32. Eaton, B., Gregory, J., Drach, B., Taylor, K. and Hankin, S. NetCDF Climate and Forecast Metadata Conventions. http://www.cgd.ucar.edu/cms/eaton/netcdf/CF-current.htm, 2000.
33. Butler, R., Engert, D., Foster, I., Kesselman, C., Tuecke, S., Volmer, J. and Welch, V. A National-Scale Authentication Infrastructure. IEEE Computer, 33 (12). 60-66. 2000.
34. Foster, I., Kesselman, C., Tsudik, G. and Tuecke, S., A Security Architecture for Computational Grids. In Proceedings of the 5th ACM Conference on Computer and Communications Security, (1998), 83-91.